

DOCUMENT RESUME

ED 455 241

SP 040 170

AUTHOR Galuzzo, Gary R.; Leali, Shirley A.; Loomis, Diane
TITLE Do We Have To Give Standardized Tests of Teacher Content Knowledge?
PUB DATE 2000-11-00
NOTE 11p.; Paper presented at the Annual Meeting of the National Council of States (21st, Miami, FL, November 17-21, 2000).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Elementary Education; Evaluation Methods; *Feedback; Higher Education; *Knowledge Base for Teaching; *Mathematics Education; *Pedagogical Content Knowledge; Performance Based Assessment; Preservice Teacher Education; Standardized Tests; Student Teachers
IDENTIFIERS Teacher Knowledge

ABSTRACT

This study examined, through a designed simulation of a performance assessment of a typical teaching act, the possibility of measuring elementary preservice teachers' mathematics content knowledge without relying solely on the candidates' performance on a standardized mathematics test. Data collection involved a series of disclosed mathematics problems designed by the National Council of Mathematics, which were used as prototypes for the National Assessment of Educational Progress (NAEP) tests. The mathematics problems were prepared for the fourth grade NAEP assessments. In this study, 24 fifth graders were required to solve six two-step problems, show their work, and explain how they solved them. The same problems were given to 69 preservice elementary teachers to examine their mathematics knowledge. They had previously taken the standardized test of mathematics skills. The preservice teachers were then given the students' responses to the problems and directed to analyze them and feedback to each fifth grader. The feedback was categorized into four types (general, specific, student performs, or teacher performs). Results indicated that the correlation between the standardized test (as the test of preservice teachers' knowledge of mathematics) and the quality of the feedback variable (the performance assessment) showed a positive but low relationship. (SM)

S. A. Leali

Do We Have to Give Standardized Tests of Teacher Content Knowledge?

Gary R. Galuzzo, George Mason University

Shirley A. Leali, Weber State University

Diane Loomis, George Mason University

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it

☐ Minor changes have been made to
improve reproduction quality

• Points of view or opinions stated in this
document do not necessarily represent
official OERI position or policy

Paper presented at the National Council of States, 21st Annual National Conference

Miami Beach, Florida

November 18, 2000

ED 455 241

The confluence of politically-charged initiatives to improve the quality of teachers entering the nation's classrooms is creating new responses to some rather long-unquestioned assumptions about what teachers should know and be able to do. From one side of the spectrum, the profession of teaching is challenged by those who believe that successful performance on standardized tests of teacher content knowledge is both necessary and sufficient for gaining access to a classroom (Educational Week, Quality Counts, 1999). From the other side of the spectrum, the profession is being forced to think about how teaching is a learned skill which goes beyond the mere possession of subject matter knowledge and which necessitates performances that have attributes of effectiveness that can only be determined via professional preparation. Languishing somewhere in the great middle are the many who readily acknowledge that both subject matter knowledge and teaching performance assessment are necessary conditions for teaching effectiveness, but who remain skeptical about the measures which both ends of the spectrum tout as essential for determining who can become a licensed teacher.

Illustrative of the trend toward teacher testing is the fact that almost all states now require teachers to pass examinations as a condition of state licensing. Indicative of the skepticism within the profession, the Public Agenda reports, teacher educators hold a certain disdain for standardized testing. As the authors of its report, Different Drummers write, "the idea that filling in a circle could measure learning seems improbably to most education professors. They don't believe standardized tests demonstrate learning nor are they willing to place high stakes on the results of such exams." (Public Agenda, 1977, p.13). It is an unsubstantiated assertion, but likely generally true, that teacher educators see the limiting weaknesses of standardized tests, how exclusive they can become, and how they've been misused over the decades. For most critics of standardized testing, the issue is not the reliability and validity of the instruments

themselves. Rather, the issue is whether what these tests reveal is what really matters. At this end there are those who believe that standardized testing of teacher content knowledge misses the point completely when it is substituted as an indication of whether a teacher can teach the discipline effectively. There is no doubt that one cannot teach well what one does not know well. The question remains, however, "Is knowledge of the discipline also a measure of teacher effectiveness in student achievement?"

At the other end are those who believe the better way to know whether a teacher is qualified to teach and possesses the likelihood of being effective is through performance assessment, direct observation, and the study of teachers' products (Diez 1993). It is argued that the assessment of performance provides a more rounded picture of what someone knows and is able to do. These types of situ, applied performance activities, or simulated assessments of teaching, allow the person being assessed to represent more fully his/her knowledge and skills.

The growing recognition that there is more to teaching than subject matter preparation has created an interest in determining new ways to assess teacher competence. The rising tide of thought on performance assessment, the successful work of the National Board for Professional Teaching Standards (NBPTS) and the promising work of the Interstate New Teacher Assistance and Support Consortium (INTASC) provide some insight into how we might more effectively assess preservice teachers' knowledge of the discipline(s) they will teach without relying completely on a score on a standardized test to make that determination.

Objective and Purpose of the Study

The objective of this study was to examine, through a designed simulation of a performance assessment of a typical teaching act, whether it is possible to measure elementary preservice teachers' mathematics content knowledge without relying solely on the candidates' performance on a paper-and-pencil mathematics test. The purpose of the study was to test whether preservice teachers' mathematics knowledge is evident in the quality of feedback they give on students' work.

This study was based in five perspectives. From the scholarship perspective, it was borne most directly from the work of Shulman (1986) and his associates who have been examining the concept of

pedagogical content knowledge in the disciplines. Shulman argued that teachers possess both knowledge of content and knowledge of pedagogy. He also argued they possess a unique blending of these two into what he labeled pedagogical content knowledge. The specific dimensions of pedagogical content knowledge have yet to be determined fully, but the work of Ball, Lampert, Clift, Grossman, McDiarmid, Wilson, and Wineberg, among many others has shed some light on the context-specific dimensions of content-related pedagogy. The present study, finding its underpinnings in this unique blend of content knowledge of preservice teachers contributes to their ability to provide feedback on students' work.

The second perspective was drawn from the work of Brophy (1981) on effective teacher praise and feedback. His review of the literature on teacher praise suggested twelve attributes or characteristics that make teacher praise more effective in promoting student achievement. These include whether the feedback was specific to the student, stated the accomplishment, was conditional only to the interaction under review, and more. Generally, Brophy's review of the literature suggests that teachers who provide specific, useful, and through feedback to the student are likely to increase student performance. In contrast, less effective praise is general, does not state the accomplishment, often is global, and has not been found to correlate with student achievement.

One dimension, we argue, of high-quality teacher feedback is knowledge of the discipline being taught. In order for the teacher to give useful and specific praise that advances student understanding, he/she must know the discipline well-enough to be specific and thorough in his/her praise. As will be seen in later in this paper, the ability to provide specific feedback was an essential component in this study.

The third perspective is based in the political reality that almost all states now require a test of teacher knowledge as a condition for gaining access to a classroom. There is a general belief that one cannot teach what does not know. In order to address this observation, state legislatures around the country are enacting laws that include teacher testing in the basic skills at a minimum. Two testing companies dominate in this area: Educational Testing Service through its PRAXIS series, and National Evaluation Systems through its customized tests for states. It cannot be escaped that some who enact these laws see teacher content knowledge and the ability to teach as being the same event. This study explores that

untested relationship.

The fourth perspective was drawn from the teacher assessment literature, probably best represented by the many studies conducted to support the work of the National Board for Professional Teaching Standards. Under contract to the National Board, those developing these assessments highlighted the problems associated with teacher assessment. Particular to these is the issue of estimating teacher knowledge without requiring practiced teachers seeking National Board Certification to take paper and pencil exams of content knowledge. The National Board has invested deeply in the design of performance assessments which capture the complexity of accomplished teaching, are professional in nature, and are respectful of the time it takes teachers to complete them. Thus, performance assessments have become the medium for determining whether a candidate earns National Board Certification.

The fifth perspective supporting this study was based in the teacher education program evaluation and research literature. In this literature, the field of teacher education has roundly been criticized for lacking a base in inquiry, both generally (Doyle, 1997; Yarger and Lee, 1997) and at the program level (Galluzzo and Craig, 1990). That is, these writers and others have argued that we know far too little about the effectiveness of a field practiced as pervasively as teaching is practiced. Faculty members know little about the effects of their programs. Institutions have made marginal commitments to systematic research and evaluation in teacher education. The present study sought to find an alternative method of assessing teacher knowledge so that the faculty would have some sense of the effects of the program on preservice teacher's skill development, and that would also be a part of the program itself and not additional work unrelated to teacher professional development.

Bringing these five perspectives together, this study was designed to examine preservice teachers' knowledge of mathematics through a performance assessment that required them to simulate teacher feedback on students' mathematical reasoning. The data would be gathered from the teacher feedback statements on students' work to determine whether the program was having its desired effect of preparing teachers for the classroom who are ready, in this instance, to provide quality feedback on students' performance in mathematics. Thus, the fundamental assumption which guided this study was that the better a teacher knows the discipline to be taught, the more likely it is that the quality of feedback will be

richer and informative toward building student understanding of that discipline.

Method

The method used to gather the data for this study included a series of disclosed mathematics problems designed by the National Council of Mathematics (NCTM), which were used as prototypes for the National Assessment of Educational Progress (NAEP) tests. The mathematics problems were prepared for the fourth-grade NAEP assessments. The items used in the present study required the pupils to solve six two-step mathematics problems, to show their work, and to write an explanation of how they solved them. These items were given to a class of twenty-four fifth-graders.

The same problems were given to sixty-nine preservice teachers in an elementary teacher education program at a western state university to examine their knowledge of mathematics. The preservice teachers had previously taken the Educational Testing Service's Pre-Professional Skills Test: Mathematics (PPST) as a state licensing requirement. The PPST "measures those mathematical skills and concepts that an educated adult might need. It focuses on the key concepts of mathematics and on the ability to solve problems and to reason in a quantitative context." (Educational Testing Service, 1999).

The preservice teachers were then given the students' responses to the mathematics problems, and were directed to analyze them for the purpose of giving feedback to each student in the fifth grade class. In this way, the researchers sought to simulate a routine in which practicing teachers engage daily: giving written feedback on student work. Through this activity, the preservice teachers were providing data to the researchers on their knowledge of math content (related to NCTM content standards) and on the pedagogical practice of giving feedback to students. The researchers analyzed the preservice teachers' quality of feedback on the math problems.

The feedback was categorized into four types drawn from Brophy's analysis of teacher praise. The praise or feedback was termed either "general" or "specific." An example of general (non-specific) praise is, "Great job! keep it up!" In contrast, preservice teachers feedback that was scored as "specific" included statements such as, "Your use of visuals helped you solve the problem" and "it is always a good idea to check your work, as you did." The remaining two categories related to the kind of follow-up performance

the teachers' feedback elicited from the student. In most cases, the teacher asked the student for continuing effort to solve the problem. For example, a teacher might direct the student's attention to part of the problem, and suggest a strategy for coming closer to solving that step, and then ask the next student to work on the problem along this line. Even when a student had gotten all parts of the problem correct, the teacher might extend the student's thinking even further by asking if the student could think of another way to solve the problem. This third category was labeled "student performs." In contrast once in awhile the teacher did the work for the student and gave the correct answer for him/her. The fourth category was labeled "teacher performs." Appendix 1 displays the categories with more examples of the feedback within each one.

The data analyzed by assigning points for the different kinds of feedback. When the teacher gave the student specific praise or feedback, the response was assigned two points. For each instance in which the teacher's feedback was one of general praise, it was assigned one point, thereby not penalizing teachers for providing some feedback, when none is always a clear possibility. Teachers' responses that furthered the student's thinking (the "student performs" category) were also assigned two points. Finally, for each instance in which the teacher performed the math problem rather than the student, no points were awarded. The total number of points for each teacher was tallied, and that became the score for each teacher for "quality of feedback" which was used for data.

Results

**** Note:** The statistical correlation is being updated to reflect a more accurate analysis that is reflective of most recent data received from on-going research.

Table 1 displays the means, standard deviations, minimum and maximum scores on each measure.

Table 1

Means and Standard Deviations of Preservice Teachers on the PPST and Quality of Feedback

Measure	PPST	Quality of Feedback
X	180.42	18.54
SD	5.05	9.52
Min	171	5
Max	189	44

In order to examine how well the performance assessment measure, entitled "quality of feedback" might also represent teacher content knowledge, as traditionally measured by the PPST, a Pearson product-moment correlation was run. The correlation between the quality of feedback and preservice teacher's math knowledge was $r=.11$, ns. The low correlation indicates that this measure of teacher feedback and the preservice teachers' knowledge of mathematics as assessed by the PPST, are not measures of the same event. Preservice teachers' quality of feedback on students' work is independent of the preservice teachers' knowledge of mathematics, and therefore cannot be considered an indication of teacher content knowledge. In other words, the PPST and this performance assessment activity are measuring two different aspects of preservice teachers' knowledge.

Discussion

This study began as an examination of the unresolved point that an aspiring teacher's knowledge of the discipline is sufficient to allow access to a classroom. The implications of that point for teacher preparation were operationalized in an hypothesis that it is possible to design teacher performance assessments that would serve as a better representation of what a preservice teacher knows and is able to do. It was hoped the performance assessment would capture not only the ability to perform the teaching skill under study, but would also indicate how well the preservice teachers "knew" the discipline they were teaching based on the quality of the feedback given to the students of the preservice teachers. It was then argued that the quality of preservice teachers' feedback on students' mathematics papers would depend to some extent on how well the teacher "knows" the subject, in this case mathematics. The correlation between the PPST, as the test of preservice teachers' knowledge of mathematics, and the quality of feedback variable (the performance assessment) shows a positive, but low relationship. While the correlation between these two measures is positive, it is too low to encourage those who question the validity of standardized testing of teachers, and the typical substituting of that score for (eventual) teaching performance, to find much hope. In other words, the ability to provide clear, specific and accurate feedback, as designed for this study, is virtually unrelated to preservice teachers' knowledge of the discipline, as measured for this study.

The reasons for the low correlation could be many. First, it could be that the measures used in this study were not sensitive enough to illuminate the problem. For example, it could be that the mathematics

measured on the PPST is at such a level that it does not capture elementary preservice teacher content knowledge well enough for research purposes. It is difficult, however, to make this argument with the PPST. The reliability and validity of tests such as the PPST are sound as the tests are prepared by companies that devote millions of dollars to research and development. To dismiss the PPST as a measure that lacks content validity is to dismiss a well-developed and well-documented test.

A second explanation for the low correlation and lack of significance may be the performance assessment task. It could be argued that the math items used in the performance assessments in this study were lacking in some way; perhaps that they didn't represent the teaching of mathematics well. Again, the items used in this study were prototypical NCTM items which had the support of the profession and therefore a measure of content validity, at the least. Additionally, the tasks performed by the preservice teachers were routine tasks any teacher performs: analyzing student work and providing feedback that encourages growth.

A third reason for the lack of a stronger correlation could be the scoring of the teacher feedback. The scoring scheme created for this study followed the spirit of the values of the research on teacher praise and feedback as summarized by Brophy. It seemed to make sense to award more points for specific praise and feedback and for teacher statements intended to extend students' thought, and fewer points for general, non-specific praise or feedback and for teachers' doing the work for the students. Other approaches to scoring schemes may reveal different results.

A fourth reason may be that the relationship between preservice teacher mathematics knowledge and the ability to use that knowledge for student feedback is not clear in the literature. It may have been the researchers' "leap of faith" that high-quality teacher feedback has high-quality teacher content knowledge as a prerequisite condition. This is a topic for future research, for the field may not know well enough the influence that teacher content knowledge exerts on the subtleties and complexities of teaching. It would seem that greater knowledge of the content being taught would allow for more specific teacher feedback to the students, just as one can speculate that those who lack content knowledge are limited in the amount of specific they can give to the students. Again, this may be a weakness in the premise of the study, but which for the most part is unexamined in the literature.

On a more positive note, however, the low correlation between preservice performance on the PPST and their quality of feedback can also be seen as reason to study further the domains of teacher knowledge and performance. An argument can be made that the low correlation also means that standardized tests do not generate enough data on which to make decisions about who can/should gain access to a classroom as a teacher. In other words, highly competent performance on a standardized test such as the PPST does not necessarily imply equally highly competent performance in providing students with accurate and helpful feedback on their performance. That finding from this exploratory study suggests that the proper assessment of preservice teachers requires both a test of teacher subject matter knowledge and a set of performance assessments to examine how well preservice teachers are prepared to perform this important dimension of teaching. If this study provided any sense of direction, it is that performance on standardized tests alone does not generalize to teaching skill. Those who question standardized testing of teachers may find some support for their argument in this analysis.

As for the issue of utility in teacher education programs, this study demonstrates the following. First, this approach to examining the pedagogical content knowledge of preservice teachers can be done at the program level. That is, preservice teachers can be asked to demonstrate their range of ability with regard to pedagogical content knowledge as part of their preparation program. Second, the resultant data will tell program faculty much more about their aspiring teachers' skill than will be determined by a standardized test of either content knowledge or pedagogical knowledge, for that matter. Finally, such types of studies, which can become part of a portfolio of program assessment tasks, serve the two purposes of advancing the field in areas of limited knowledge and increasing institutional accountability at a time when that is a high priority in the policy arena.

The field of teacher education is not without perspectives on learning to teach that can challenge the prevailing belief that teacher content knowledge is both necessary and sufficient for successful teaching. This study sought to examine the widely-supported "alignment of standards" by studying preservice teacher content knowledge through their feedback on students' work. The simulation was tied to national math content standards. The purpose was to begin determining whether we can model new forms of preservice teacher assessment and eventually use such assessments of teacher content knowledge in action as a method for determining licensure and program effectiveness. In that regard, this study may

inform the discussion of teacher education and teacher assessment.